IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

UTILITY PATENT APPLICATION FOR:

CHIP MULTIPROCESSOR WITH MULTIPLE OPERATING SYSTEMS

Inventors:

Stephen E. RICHARDSON
668 Vera Cruz Avenue, Los Altos, CA 94022

Gary VONDRAN
1400 Magnolia Avenue, San Carlos, CA 94070

Stuart SIU
20943 Elbridge Court, Castro Valley, CA 94552

Paul KELTCHER
878 Hollenbeck Avenue, Sunnyvale, CA 94087

Shankar VENKATARAMAN
1111 Derbyshire Drive, Cupertino, CA 95014

Padmanabha VENKITAKRISHNAN
961 Bermuda Court, Sunnyvale, CA 94086

Joseph KU
4102 Amaranta Avenue, Palo Alto, CA 94306

CHIP MULTIPROCESSOR WITH MULTIPLE OPERATING SYSTEMS

FIELD OF THE INVENTION

The present invention is generally related to a computer chip architecture having multiple

5   processors on a single die. More particularly, the present invention is related to a

multiprocessing chip utilizing multiple operating systems.

BACKGROUND OF THE INVENTION

Existing internet data centers (IDCs) pack hundreds of processors (e.g., servers and the

10   like) in a single building for processing a large volume of data transactions. Generally, the

compute density or number of nodes per volume defines the efficiency of the IDC. The compute

density effects the amortization of the high cost of the IDC infrastructure (e.g., networking,

power, cooling, maintenance, reliability, and availability support). Typically, the greater the

compute density, the better the IDC will be able to amortize the high cost of IDC infrastructure.

15   Accordingly, a large compute density may be preferred. However, space may be unavailable or

costly for locating a large number of processors necessary for maintaining a large compute

density.

To provide increased compute density, multiprocessing schemes that utilize multiple

processors have been developed. One conventional multiprocessing scheme (shown in Fig. 1)

20   includes a computer system 100 having multiple processors 10-40, each on a separate die (i.e.,

chip) 50-80, and connected to a single operating system 90 stored in a memory 95. The system

100 may conserve space if the system is provided in a single housing.

A second conventional multiprocessing scheme (shown in Fig. 2) includes a computer system 200 having a chip multiprocessor 295. The chip multiprocessor includes multiple processors 210-240 on a single die 290. Similar to the processors 10-40 in system 100, the processors 210-240 are connected to a single operating system 250 stored in a memory 260. The processors 210-240 may communicate with the memory 260 via a bus 270. System 200 conserves space by providing multiple processors on a single die. However, the systems 100 and 200 suffer from well known scalability problems.

Schemes that have placed multiple processors on a single chip typically utilize a single operating system for tying all the processors together. A well known limitation of this scheme and other multiprocessing schemes utilizing a single operating system is that an operating system does not scale well to large numbers of processors. That is, as the number of processors managed by a single operating system increases, the efficiency of the operating system goes down dramatically. For example, an operating system typically includes internal data structures that may be limited in the number of processors that can be supported, and limited bandwidth on a bus may slow transactions. Thus, scaling becomes impractical above some small number (e.g., currently about four to at most 64 processors, depending on the operating system in question).

Bugnion et al., in U.S. Pat. No. 6,075,938, discloses using a cache coherent non-uniform memory architecture (CC-NUMA) that supports multiple processors executing multiple operating systems. However, Bugnion et al. discloses multiple virtual processors, implemented in software on a single physical processor. This architecture fails to provide multiple physical processors, implemented in hardware, on a single die. Accordingly, this architecture suffers a performance penalty, because the single physical processor must task switch among multiple virtual processors (only one virtual processor can be running on the physical processor at any

given time). Moreover, if this architecture were to support multiple physical processors, it would

need space for providing multiple dies, and processing speed would consequently be sacrificed

due to the input/output procedures needed to communicate among the multiple separate

processors and the memory.

5

SUMMARY OF THE INVENTION

An aspect of the present invention is to provide a multiprocessing system including

multiple processors mounted on a single die. The multiple processors are connected to a

memory storing multiple operating systems. Each of the multiple processors may execute one of

10   the multiple operating systems.

Another aspect of the present invention is to provide a multiprocessing system including

a plurality of processor groups mounted on a single die. The processor groups are connected to a

memory storing multiple operating systems. Each of the processor groups may execute one of

the multiple operating systems. The processor group may include one or more processors

15   mounted on the die.

Certain embodiments of the present invention are capable of achieving certain

advantages, including some or all of the following: mounting multiple processors on a single die

reduces the cabling problem inherent in connecting multiple processors on separate dies in

separate housings; mounting multiple processors on a single die reduces the latency required for

20   communication among the processors and improves the efficiency of message passing,

potentially enabling a whole new class of applications (e.g., data mining) to run on such a

multiprocessing system; mounting multiple processors on a single die reduces chip-to-chip

communication costs and leads to further power efficiency; and increased scalability for

multiprocessing.

Those skilled in the art will appreciate these and other advantages and benefits of various

embodiments of the invention upon reading the following detailed description of a preferred

5   embodiment with reference to the below-listed drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention is illustrated by way of example and not limitation in the

accompanying figures in which like numeral references refer to like elements, and wherein:

10   Fig. 1 illustrates a conventional multiprocessing scheme including multiple processors,

each on a separate die;

Fig. 2 illustrates a conventional multiprocessing scheme including multiple processors on

a single die;

Fig. 3 illustrates an embodiment of a multiprocessing scheme employing the principles of

15   the present invention; and

Fig. 4 illustrates another embodiment of a multiprocessing scheme employing the

principles of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

20   In the following detailed description, numerous specific details are set forth in order to

provide a thorough understanding of the present invention. However, it will be apparent to one

of ordinary skill in the art that these specific details need not be used to practice the present

invention. In other instances, well known structures, interfaces, and processes have not been shown in detail in order not to unnecessarily obscure the present invention.

Fig. 3 illustrates an embodiment including a computer system 300 employing the principles of the present invention. System 300 includes a chip multiprocessor 350 having

5   multiple processors 305-320 mounted on a single die 360. The processors 305-320 function with operating systems 325-340 respectively, as illustrated by connections 345-348. The operating systems 325-340 are stored in a memory 365. During operation, each processor 305-320 may access a respective operating system 325-340 by communicating with the memory 365, for example, via a bus 370. Alternatively, each processor may be directly connected to the memory

10   365 without using the bus 370. The memory 365 may include one or more of the following: SRAM and/or DRAM on the same chip as one or more processors; SRAM and/or DRAM on separate chips connected to one or more processors; magnetic media, such as tape or disk; optical media, such as CD-ROM; and the like.

Processors 305-320 may be configured, such that each processor executes its own operating system. Multiple processors (e.g., multiple processors in a processor group) may also

15   be configured to execute a single operating system. Fig. 4 illustrates a second embodiment including a computer system 400 employing the principles of the present invention. System 400 includes a chip multiprocessor 450 having multiple processors 405-420 mounted on a single die 460. The processors 405-420 are divided into two processor groups 425 and 430 having

20   processors 405 and 410 in the processor group 425 and processors 415 and 420 in the processor group 430. The processor group 425 executes an operating system 435 stored in a memory 465 (as illustrated by a connection 475), and the processor group 430 executes an operating system 440 stored in the memory 465 (as illustrated by a connection 480). During operation, each

processor group 425-430 may access a respective operating system 435-440 by communicating with the memory 465, for example, via a bus 470. Alternatively, each processor and/or processor group may be directly connected to the memory 465 without using the bus 470.

The operating systems shown in Figs. 3-4 and described above may include conventional

5   operating systems, such as WINDOWS NT, UNIX and the like, and the processors in systems 300 and 400 may include conventional processors. Each processor may be capable of executing a single operating system, or capable of simultaneously executing multiple operating systems, for example, by context switching, which may include rapidly switching between multiple operating systems. Systems 300 and 400 may be operable to execute a variety of applications, such as web

10   service, database service and the like. Four processors are shown in Figs. 3-4 for illustration purposes, and it will be apparent to one of ordinary skill in the art that more or less processors and/or processor groups may be included on the dies 360 and 460. Additionally, as is known in the art, the processors in systems 300 and 400 may access one or more caches (not shown).

It will be apparent to one of ordinary skill in the art that a system employing the

15   principles of the present invention may be operable to support both processor groups and processors on a single die, such that each processor group and processor not within a processor group executes a distinct operating system.

Having each processor (or processor group) executing its own independent operating system minimizes scaling problems realized when utilizing a single operating system with

20   multiple processors. For example, one hundred processors independently executing one hundred operating systems may be no more of a problem than one processor executing one operating system. The embodiments described above are operable to provide this type of scaling on a single chip.

While this invention has been described in conjunction with the specific embodiments thereof, it is evident that many alternatives, modifications and variations will be apparent to those skilled in the art. There are changes that may be made without departing from the spirit and scope of the invention.

5